

DATA SECURITY FOR DIGITAL DATA STORAGE

a1

Background of the Invention5 Field of the Invention

The invention relates to methods and apparatus for providing security for digital data stored on data storage media such as magnetic and optical disks and tapes.

Description of the Related Art

10 Over the past several years, personal computing systems have become more powerful, smaller, and less expensive. As this has occurred, more and more computing applications are performed on personal computer platforms. Local and wide area networks of personal computers are now often used in corporate and business applications instead of the large mainframes used for the same applications in the past. A further result of the increases in performance and decreases in price of personal
15 computers is a dramatic increase in personal computer use for household financial and other sensitive and preferably confidential information.

The use of personal computers in these applications raises data security and privacy issues which have thus far been insufficiently resolved. Laptop and other personal computers, as well as the removable data storage media used in them are
20 transported, mislaid, lost, and sometimes stolen. Consequently, security and privacy issues which were not present when computers and their data storage media were generally fixed have now become prominent. Administrators of computer resources in the business environment must remain aware of the location of portable computing devices as well as the nature of the programs and data stored on them. For home users,
25 concerns arise if credit card, social security, or bank account numbers are present on fixed or removable media which may be lost or stolen.

To resolve a few of these concerns, some programs allow the user to password protect documents or files, thereby preventing access to the data in the file unless the password is known. This provides limited security, however, since these schemes are
30 easy to defeat with widely available password extraction programs. Furthermore, although the act of opening the file may be restricted in the relevant application

program, the data itself resides on the media in raw form, and may still be extracted by a trained computer user.

5 Systems have also been proposed which perform encryption on data and application programs stored on tape and disk. These systems provide improved security over the password protection described above. As one example, a system disclosed in U.S. Patent No. 5,325,430 to Smyth et al. (incorporated herein by reference in its entirety) includes a security module attached to a personal computer which performs data and application program encryption. The security module communicates with a removable smart card assigned to a given user which contains encryption keys used by the security module. Although the security provided by this system is adequate for many applications, the circuitry used to implement the system is complex, and administration of the system for producing and assigning keys and smart cards is time consuming and expensive.

15 Another system for encrypting files is disclosed in U.S. Patent No. 5,235,641 to Nozawa et al., the disclosure of which is also incorporated herein by reference in its entirety. In this system, data stored to a magnetic tape is encrypted by a cryptographic adapter which is located in the data path between a host processor and a tape drive. In this system, the host processor generates cryptographic keys which are stored on the tape itself. This requires additional logic to encrypt the keys as well as the data, and consequently, this system requires relatively complex circuitry, and leaves the key potentially recoverable from the tape itself if the key encryption scheme is broken.

20 Thus, existing encryption systems for personal and portable computers have serious drawbacks, and have not been widely implemented. In particular, a system which is useful for an individual personal computer user has not been heretofore provided. Such a system should provide data security with flexibility and without expensive administration or implementation.

Summary of the Invention

25 In a first embodiment, the invention includes a computing apparatus comprising a digital data storage device and a logic circuit configured to receive digital data and to forward the digital data to the digital data storage device in an encrypted form. The computing apparatus may also comprise a non-volatile memory location which stores an

identification code, as well as a second memory location coupled to the logic circuit which stores an encryption key accessed by the logic circuit derived at least in part from the identification code.

5 Another embodiment of the invention includes a circuit for encrypting data in a computing system. The circuit may comprise a first memory location storing an identification code, and a logic circuit comprising a second memory location and an encryption engine. The logic circuit may be configured to receive the identification code and to store a key for use by the encryption engine derived at least in part from the identification code in the second memory location.

10 The invention also comprises a computer having a plurality of data storage media drives and a data path connected between a source of data and the data storage media drives. The computer may further include a logic circuit coupled to the data path which is configurable to enable encrypting of data being routed to a selectable subset of the plurality of data storage media drives.

15 Brief Description of the Drawings

FIG. 1 is a block diagram of a data storage system incorporating an embodiment of the invention.

FIG. 2 is a flow chart illustrating acts performed during key generation in an embodiment of the invention.

20 FIG. 3 is a block diagram illustrating an encrypting data path passing from a host processor to data storage devices, in accordance with one embodiment of the invention.

FIG. 4 is a flow chart illustrating acts performed during key generation in another embodiment of the invention.

25 Detailed Description of the Invention

Embodiments of the invention will now be described with reference to the accompanying Figures, wherein like numerals refer to like elements throughout. The terminology used in the description presented herein is not intended to be interpreted in any limited or restrictive manner, simply because it is being utilized in conjunction with a detailed description of certain specific embodiments of the invention. Furthermore,
30 embodiments of the invention may include several novel features, no single one of

which is solely responsible for its desirable attributes or which is essential to practicing the invention herein described.

Referring now to Figure 1, a data storage system is illustrated which incorporates aspects of the invention. The system includes encryption/decryption logic 10 that is connected to receive digital data from a data bus 12. The encryption/decryption logic 10 is configured to forward data received from the data bus 12 to data storage devices 14 in an encrypted form. The data or information transferred between the data bus 12 and the data storage devices may comprise application programs themselves, data used by application programs, or any other information that the host computing system stores to the data storage devices 14 of the system. As will be further explained below with reference to Figures 2 and 3, the encryption/decryption logic may in some embodiments be configurable to perform the encryption and decryption on a selectable subset of the data storage devices if desired by a user of the system.

The algorithm used to perform the encryption may comprise any known encryption algorithm, and many different alternatives will be well known to those of skill in the art. In many applications, the encryption and decryption process will be defined in part by a key 16 which is utilized by the encryption/decryption logic 10 to perform the data manipulation which results in data encryption and decryption. In some systems, the same key is used for both the encryption and decryption processes, but in others, the key 16 may comprise a pair of keys, wherein one is used for encryption, and the other for decryption. Given the variety of encryption and decryption schemes which have been and are currently being developed, the use of the word "key" is intended to encompass any piece of information, data, parameter, definition, configuration of logic circuitry, or other entity or circuit arrangement which serves at least in part to configure the encryption/decryption logic, or to otherwise in any way partly or wholly define the data encryption process which is performed by the encryption/decryption logic 10.

Also provided in the system of Figure 1 is a non-volatile memory location 18. As is well known in the art, a non-volatile memory has the property that the data or information stored in it remains when the host system is powered down. Non-volatile memory may comprise battery backed up RAM, EPROM, EEPROM, or ROM memory circuitry, for example. In the application of Figure 1, this memory location 18 may

advantageously store an identification code. The stored identification code may be used to derive, at least in part, the key 16 which is used in the encryption process. This derivation may involve simply making the key the identification code itself, or may alternatively involve a logical or mathematical manipulation or transformation of the identification code to produce the key. In some embodiments, as will be further explained below, the key 16 may be derived in part from the identification code stored in the non-volatile memory and in part from a password or other piece of information entered by a user of the computing system.

The system of Figure 1 includes many advantages over prior art data encryption schemes and is especially applicable to individual personal computer and laptop computer users. In some embodiments, the circuitry of Figure 1 may be incorporated into, for example, a laptop computer which is sold to an individual for household and/or business use. In most of these situations, the purchased computer will not be a member of a group of computers which is controlled or overseen by a system administrator that will create and assign encryption keys, smart cards, etc. Rather, the laptop will be simply used as is, for both personal and business use by a user who is generally unfamiliar with data security techniques or procedures.

In these embodiments, the identification code may comprise a multi-bit data word which is associated with the individual laptop being used. When stored in a non-erasable memory element such as ROM or EPROM, the identification code may be substantially permanently associated with the individual laptop being used. It will be appreciated that data security in these environments is enhanced if different laptops do not typically share a common identification code. When this is true, the key 16 derived from the identification code will be different in different laptops produced by the laptop manufacturer. It will therefore also be appreciated that the data stored on the data storage devices 14 will be encrypted differently by different laptops. Thus, a removable media such as a floppy disk, tape, or writeable CD will not be useable on any computer except the one that originally stored the data. Thus, a level of security is provided for removable media which may be lost, mislaid, or stolen.

It will also be appreciated that this level of security is provided without any intervention by the user or a system administrator. Key generation and data encryption is

automatic and transparent. In addition, this data security scheme is easily implemented in the large scale production of laptops and other personal computers. Machine specific data encryption may be provided with the simple provision of non-volatile storage of information which defines the data encryption process performed. This information may advantageously be substantially uniquely associated with the host computing logic or host computer. This may be ensured by using some form of sequential numbering scheme for the identification code, or alternatively a random or pseudo-random numbering scheme with a low probability of producing two identical identification codes for different laptops. However, it may be noted that it is not necessary to absolutely guarantee that each laptop have a uniquely defined encryption process. The desirable feature is that there be a relatively low probability that lost or stolen media will be readable in some other laptop or personal computer available to someone who has found or has stolen the media elements. Therefore, duplicate identification codes and keys defining identical encryption processes may be provided within a given set of computers while still maintaining a useful level of security. Thus, the association between identification codes and their respective host computers need only be substantially unique such that a reasonable level of security is created.

Figure 2 illustrates a method of key generation and data encryption according to one embodiment of the invention. It will be appreciated that the method shown in Figure 2 may, in one embodiment, be implemented on hardware illustrated in Figure 1.

The method begins at a start state 22, and moves from there to step 24, where an identification code is retrieved. The identification code may be stored in a non-volatile memory, and may in addition be substantially uniquely associated with specific host computer hardware.

The system then moves to decision state 26, where it is decided whether or not some user input should be utilized in the process of encryption key generation. If not, the method moves directly to step 28, where an encryption key is generated using the identification code retrieved at step 24. If user input is to be used in key generation, the method moves from step 26 to step 30, where the user input is accepted by the system. The user input may, for example, comprise an alphanumeric code which is typed into the computer keyboard by the user in response to a system prompt. Following this, the

method moves to step 28, where the key is generated using both the identification code and the user input. The user input from step 30 may be an alphanumeric sequence which is converted to a multi-bit word (for example, to ASCII code). This word may be combined with the identification code in many ways, including concatenation as one simple example, or other more complicated logical or mathematical manipulations may be used.

Following key generation at step 28, the key is used to encrypt and decrypt data that is stored to and retrieved from a data storage device at step 32. In the personal computer or laptop computer context, the series of steps leading to and including key generation may be performed during the boot operation prior to any accesses to encrypted data storage devices. In these embodiments, all data, programs, etc. stored on the data storage devices are encrypted with the same key, a key which may require some user input to generate as described above. The computer may be either factory configured or user configured to require or not require user input for key generation.

The addition of user input to key generation provides an enhancement to data security beyond that provided when only the identification code is used to derive an encryption key. This is because if the entire computer is lost or stolen, when the computer is turned on only the computer owner will know what code or password to input in order to generate the proper key at step 28 of Figure 2. Thus, access to encrypted programs and data is effectively prevented even with the original computer the hands of an unauthorized user.

An embodiment of the invention is also illustrated in Figure 3 which may be used to implement the process described above with reference to Figure 2. In this Figure, a computer system is shown having a host processor 36, which may, for example, comprise a member of the Pentium® family of processors such as the Pentium, Pentium Pro, or Pentium II. Although industry standard PC architecture is used as an illustrative example in this Figure, it will be appreciated that many computer designs may be implemented using the principles illustrated herein. Also provided as part of the computer system of Figure 3 are a plurality of data storage devices, including hard disk drives 38, 40, a floppy disk drive 42 and a CD drive 46, which may be of a writeable type.

The processor 36 interfaces with a host bus 44 which also interfaces with a bridge

169920-0369 circuit 46. The bridge circuit 46 routes data from the host bus 44 to a PCI bus 48. The PCI bus 48 provides a data source to a logic circuit 50 which is provided in the data path between the PCI bus 48 and an IDE bus 52 and floppy drive control bus 54 which interface directly with the respective data storage devices 38, 40, 42, 44, and 46. The PCI
5 bus 48 may also receive data from I/O devices 56 via a PCI to ISA bridge circuit 58.

The logic circuit 50 advantageously includes an encryption engine 60 which operates to encrypt data routed to one or more of the data storage devices 38, 40, 42, 46 and to decrypt data routed from one or more of the data storage devices 38, 40, 42, 46. The logic circuit 50 will also generally include input and output bridge circuitry 51 to
10 buffer data and convert the data transfer protocol from the PCI format bus 48 to the busses 52, 54, which interface directly with the data storage devices 38, 40, 42, 46.

The encryption engine operates under the control of control logic 62. The control logic, in turn, receives information for controlling the encryption engine from three sources. The first is a memory location which stores a hardware identifier 64. As
15 described above, this hardware identifier 64 may be substantially uniquely associated with the computer hardware. The memory may comprise a non-volatile writeable or read only memory to help ensure essentially permanent storage of the hardware identifier 64. As is also described above, the hardware identifier 64 stored in the memory may be used by the control logic 62 (or alternatively the processor 36) to at least in part derive a key for
20 encryption and decryption of data to and from the data storage devices 38, 40, 42, 46. The control logic may also accept user input as described above to be used as additional information for key derivation.

This generated key may be stored in a key register 66 which also is coupled to the control logic 62. Prior to data being stored or retrieved from the data storage devices 38,
25 40, 42, 46, the key may be retrieved from the key register 66 for use by the encryption engine 60 during the encryption and decryption processes.

A configuration register 70 may also be coupled to the control logic 62. The content of the configuration register 70 may advantageously be user defined, and may include bits that determine which of the data storage devices 38, 40, 42, 46 have data
30 encrypted before storage to the media, and which have data decrypted when data is retrieved from the media. This feature provides significant flexibility to the user. A user

may, for example, want to encrypt some, but not all, data stored onto a floppy disk with the floppy drive 42. It may also be advantageous to have at least one hard drive 38 or 40, which contains DOS, Windows (TM), Unix (TM) or other operating system software, to remain unencrypted.

5 The configuration register may also contain bits which determine whether or not user input should be incorporated into the key being used to perform the encryption and decryption. In some embodiments, a different key may be stored for different drives. In this case, some of the keys may be generated with user input, and some without.

10 One advantageous aspect of the encryption system described herein is that it may be created with relatively minor modifications to currently existing integrated circuits. PCI to ISA and PCI to IDE bridges are well known and understood, and are commercially available from, for example, Intel Corporation. In one embodiment, therefore, an encryption engine, control logic, a key register, and a configuration register may be
15 incorporated into an existing bridge integrated circuit design to produce a portion of the logic circuit 50. Furthermore, individual EPROM, EEPROM and ROM memories which include pre-programmed identification codes are available commercially from Dallas Semiconductor of Dallas Texas as part numbers DS2401 and DS2430 for example. These devices include a unique 48 bit serial number in a ROM storage location which may be
20 utilized as the memory location which stores the hardware identifier 64. These memory chips are available with a serial I/O interface for reading the identification code and any other stored data. In this embodiment, therefore, a bridge integrated circuit which includes the encryption logic may interface over a serial bus to a separate memory integrated circuit which stores the hardware identifier.

25 Figure 4 illustrates a method of key generation and verification which may be implemented with the system illustrated in Figure 3. In the method of Figure 4, the control logic 62 (Figure 3) may be utilized to perform key generation and verification without intervention by the processor 36 (Figure 3). The method begins at a start state 76. Following this start block 76, the system retrieves the hardware identifier 64 (Figure 3) from the non-volatile memory location where it is stored. This retrieval process may
30 involve the sequential retrieval of a set of data words from the memory as illustrated by the loop defined by blocks 78, 80, and 82. Thus, at block 78, the control logic 62 may

output an initial address to the non-volatile memory to retrieve a first data word comprising a portion of the hardware identifier code 64. The address may then be incremented at block 80. If, at decision block 82, it is determined that the entire code has not yet been retrieved, the system loops back to block 78 and outputs the incremented address to the non-volatile memory to retrieve another segment of the code.

Once the entire code has been retrieved, at block 84 the control logic 62 may then generate and verify the key. As mentioned above, the process of key generation may involve merely storing a concatenation of the data words retrieved at steps 78-82 in the key register 66. This could occur during the retrieval process, or afterwards. Alternatively, mathematical or logical manipulations may be performed on the retrieved data words prior to their storage into the key register. Key verification may also be performed in a number of ways known to those of skill in the art. For example, a checksum or CRC field may be provided in the configuration register 70 or control logic 62. If no user input is utilized in key generation, this field may be generated during an initialization sequence performed during the manufacture of either the logic circuit 50 or a computer system that the logic circuit 50 is incorporated into. If user input is utilized in key generation, this CRC or checksum field may be generated during a password initialization routine when the password to be utilized in key generation is initially entered by the user.

Following key generation and verification, the system moves to a decision state 86, where the result of the key verification of block 84 is checked. If the key is verified as good, the system moves to block 88, and the key is used to encrypt and decrypt data during data storage and retrieval operations. There are several reasons why key verification might fail. An error in reading the hardware identifier may cause faulty key generation. Tampering with the logic circuit 50 may also result in incorrect key generation. Additionally, key verification may fail because required operator input to be used in key generation has not yet been entered by a user. Thus, a failure of key verification may force user input. This is illustrated in Figure 4 by the fact that if, at decision state 86, the key has not been verified as good, the system moves to a another decision state 90. At decision state 90, the system determines whether or not user input should be accepted and used in the key generation process. If the system determines that

operator input should be accepted, the system moves to block 92, where the input is read. The system then loops back to block 84, where the key is generated using both the operator input and the retrieved identification code, and is again verified against the stored CRC or checksum field. If the operator input was the correct password, the key will be
5 verified as good at the next iteration of decision block 86, and at block 88, the key will be used to encrypt and decrypt data as described above.

If, however, the operator input was incorrect, the key verification process will fail, and the system will again move to decision state 90, where the system again determines whether or not user input should be accepted. It will be appreciated that the user may be
10 given two or more attempts to successfully input the proper password. Thus, the system may loop back to blocks 92 and 84 a plurality of times, any one of which may result in correct password entry and normal data encryption and decryption at block 88.

After a selected number of iterations of incorrect password entry, the system may decide at state 90 to refuse to accept further operator input for key generation. In this
15 event, the system moves to block 94 where the key error is flagged by the system. System response to the error flag may vary widely. The system may indicate to the user that the password entries are incorrect. The system may even be programmed to destroy the content of encrypted drives in the event the key verification process fails, or fails for a selected number of consecutive verification attempts.

20 The encryption system described thus provides data security to personal and laptop computer users in a transparent manner without requiring time consuming and expensive system administration or complex and expensive hardware. The system is especially adapted to individual users, and the high volume production of computers for these users.

The foregoing description details certain embodiments of the invention. It will be
25 appreciated, however, that no matter how detailed the foregoing appears in text, the invention can be practiced in many ways. As is also stated above, it should be noted that the use of particular terminology when describing certain features or aspects of the invention should not be taken to imply that the terminology is being re-defined herein to be restricted to including any specific characteristics of the features or aspects of the
30 invention with which that terminology is associated. The scope of the invention should therefore be construed in accordance with the appended claims and any equivalents

thereof.

THIS PAGE BLANK (USPTO)

092748-039
069260-284260